

IN THE CLAIMS

1. (Currently Amended) A method comprising:
using one or more computer processors, extracting Uniform Resource Locators (URLs)
from electronic communication; and
analyzing the URLs extracted to determine whether the electronic communication is of a
first predetermined category, said analyzing comprising generating one or more signatures using
a length of the electronic communication and the URLs extracted.
2. (Original) The method of claim 1, wherein extracting the URLs comprises extracting at
least one of a hostname, a domain name, a subsection of a domain relative link, and an Internet
Protocol (IP) address from the electronic communication.
3. (Original) The method of claim 1, further comprising performing a predetermined
operation on the electronic communication if the electronic communication is determined to be
of the first predetermined category.
4. (Currently Amended) The method of claim 1, wherein analyzing the URLs comprises:
selecting one or more of the one or more signatures generated; and
comparing the selected signatures against a plurality of predetermined signatures
generated from a plurality of known electronic communications of the first predetermined
category, wherein each predetermined signature of the plurality of predetermined signatures has
a signature based upon an electronic communication length and extracted URLs.
5. (Currently Amended) The method of claim 1, wherein generating the one or more
signatures further comprises:
computing a first hash based on the length of the electronic communication;
computing a second hash based on the URL[['']]s extracted; and
generating a signature by concatenating the first hash to the second hash.

6. (Previously Presented) The method of claim 4, wherein generating the one or more signatures further comprises using at least one of the extracted URLs as at least one of the one or more signatures.

7. (Original) The method of claim 4, wherein generating the one or more signatures further comprises generating the one or more signatures based on at least one of a protocol, a hostname, a domain name, a subsection of a domain relative link, and an Internet Protocol (IP) address from the electronic communication.

8. (Original) The method of claim 4, further comprising classifying the electronic communication to be of the first predetermined category if one of the selected signatures matches one of the plurality of predetermined signatures.

9. (Original) The method of claim 4, wherein the plurality of predetermined signatures is derived from a plurality of electronic documents reported via a collaborative submission mechanism.

10. Canceled.

11. Canceled.

12. (Currently Amended) ~~The computer-readable storage medium of claim 11,~~ A non-transitory computer-readable storage medium that provides instructions that, if executed by a processor, will cause the processor to perform operations comprising:

generating one or more signatures of electronic communication using a length of the electronic communication and Uniform Resource Locators (URLs) in the electronic communication; and

determining whether the electronic communication is of a first predetermined category using the one or more signatures generated, wherein determining includes:

selecting one or more of the one or more signatures generated based on a plurality of predetermined criteria, wherein selecting one or more of the one or more signatures generated

comprises selecting a signature if the signature represents a domain that was registered within a predetermined period of time;

comparing the selected signatures against a plurality of predetermined signatures;

and

classifying the electronic communication to be of the first predetermined category if one of the selected signatures matches one of the plurality of predetermined signatures.

13. (Currently Amended) The non-transitory computer-readable storage medium of claim [[11]] 12, wherein selecting one or more of the one or more signatures generated comprises selecting signatures representing one or more of a protocol, a hostname, a domain name, and a subsection of a domain relative link having a predetermined string of letters.

14. (Currently Amended) The non-transitory computer-readable storage medium of claim [[10]] 12, wherein the operations further comprise extracting the URLs from the electronic communication.

15. (Currently Amended) A system comprising:
a plurality of databases to store a plurality of predetermined signatures of a plurality of known electronic communications of a first predetermined category; and
a server, coupled to the plurality of databases, including:
a memory device to store a plurality of instructions; and
a processor, coupled to the memory device, to retrieve the plurality of instructions from the memory device and to perform operations in response to the plurality of instructions, the operations comprising:
extracting Uniform Resource Locators (URLs) from electronic communication generating one or more signatures using a length of the electronic communication and the URLs extracted; and
comparing one or more of the one or more signatures generated against the plurality of predetermined signatures stored in the plurality of databases to determine whether the electronic communication is of the first predetermined category.

16. (Original) The system of claim 15, wherein the URLs comprises at least one of a hostname, a domain name, a subsection of a domain relative link, and an Internet Protocol (IP) address.
17. (Original) The system of claim 15, wherein the operations further comprise selecting the one or more of the plurality of signatures based on a plurality of predetermined criteria.
18. (Original) The system of claim 15, wherein the operations further comprise performing a predetermined operation on the electronic communication if the electronic communication is determined to be of the first predetermined category.
19. (Original) The system of claim 15, further comprising a database, coupled to the server, to store a plurality of reports from which the plurality of predetermined signatures are generated.
20. (Original) The system of claim 15, wherein the plurality of databases are in a remote location from the server.